A Frequency-domain test for long range dependence.

G. Gromykov[1] M. Ould Haye[1] A. Philippe[2*]

[1]School of Mathematics and Statistics.
Carleton University, 1125 Colonel By Dr. Ottawa, ON, Canada, K1S 5B6
[2] Laboratoire de Mathématiques Jean Leray et ANJA INRIA - Rennes,
2 rue de la Houssiniere, Université de Nantes, 44 322 Nantes France.

Abstract

A new frequency-domain test statistic is introduced to test for short memory versus long memory. We provide its asymptotic distribution under the null hypothesis and show that it is consistent under any long memory alternative. Some simulation studies show that this test is more robust than various standard tests in terms of empirical size when the normality of observed process is lost.

Keywords : Long memory, dependence, time series, limit theorem, hypothesis test.

1 Introduction

Let us consider a stationary process (X_t) with a spectral density of some semi-parametric form:

$$f(\lambda) = |\lambda|^{-2d} g(\lambda) \tag{1}$$

with $0 \leq d < 1/2$ and g is an even, positive, continuous function on $[-\pi, \pi]$. We say that the process (X_t) is of short memory if $d = 0$ (i.e. the spectral density f is continuous over $[-\pi, \pi]$), and we say that the process is of long memory if $0 < d < 1/2$ (i.e. the spectral density is unbounded at zero). In time-domain terms, short memory means that the covariance function is summable (i.e. covariance function $\gamma(h) = \mathrm{Cov}(X_i, X_{i+h})$ goes to zero fast as the lag h increases) and long memory means that the covariance function is not summable (i.e. the covariance function goes to zero slowly as the lag increases). Literature is full of real world examples of both types of processes. Under mild conditions, the switch between time-domain and frequency-domain is easily made via the two representations

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma(\lambda) = \frac{1}{2\pi} \left[\gamma(0) + 2 \sum_{h=1}^{\infty} \cos(\lambda h) \gamma(h) \right],$$

and reciprocally

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda = \int_{-\pi}^{\pi} \cos(\lambda h) f(\lambda) d\lambda.$$

*corresponding author : anne.philippe@univ-nantes.fr

Since the so-called re-scaled range statistic R/S was introduced by Hurst (1951), many authors have come with several time-domain statistics for testing long memory, such as modified R/S, KPSS (Kwiatkowsk-Phillips-Schmidt-Shin), rescaled variance statistic V/S (See Mandelbrot (1969), Kwiatkowski *et al.* (1992), Lee and Schmidt (1996), Giraitis *et al.* (2003), Giraitis *et al.* (2006)). See Beran et al. (2013) for a full description of all of these statistics. According to the simulation results provided in Giraitis *et al.* (2003), V/S statistic produces testing procedure with better empirical size and power as compared to the other time domain statistics. A detailed description of V/S statistic is given in Section 3. One common drawback in these statistics is the difficulty to calibrate the window parameters in order to ensure the right empirical size (in particular for the small sample) which comes with the heavy price of a very weak power. Also, evaluation of their limiting distributions often involves approximating stochastic integrals. Time-domain based tests often suffer from the so-called high empirical size (i.e. the ratio of wrongly rejecting the null hypothesis is much bigger than the nominal level α). Here we address this problem by introducing a very promising frequency-domain statistic, for testing long memory versus short memory. This turns out to be very convenient and yields very simple limiting distributions. Our test statistic can be viewed as a contribution to this approach initiated by Lobato and Robinson (1998 (see Section 3 for a detailed description) and more recently by Bailey and Giraitis (2016) in testing unit root. However, our method bears many differences as will be explained later. The main interest of our approach is to preserve a good empirical size even when the underlying distribution is not Gaussian. From a sample $X_1, \ldots, X_n$ of the process (X_t), we are interested in building a testing procedure to discriminate between short and long range dependence. The proposed statistic is inspired from the well-known fact (see for example Giraitis *et al.* Theorem 5.3.1 (2012) or Moulines and Soulier 2003, section 3.1) that, for a large class of weakly dependent stationary processes, at any Fourier frequency, the normalized periodogram is asymptotically exponentially distributed. That is, for any fixed j, with $\lambda_j = 2\pi j/n$,

$$\frac{I_n(\lambda_j)}{f(\lambda_j)} \xrightarrow{d} E, \qquad \text{as } n \to \infty \tag{2}$$

where E has exponential distribution with mean 1, and where I_n is the periodogram

$$I_n(\lambda) = \frac{1}{2\pi n} \left| \sum_{j=1}^{n} X_j e^{ij\lambda} \right|^2 .$$

Moreover, for different Fourier frequencies, the variables on the left hand side of (2) are asymptotically independent (see same references above). Since under short memory, the periodogram is an asymptotically unbiased estimator of the spectral density f, we estimate $f(\lambda_j)$ by a sample average of periodograms. The procedure is as follows: we split our initial sample $X_1, \ldots, X_n$ into m blocks (or epochs), each of size ℓ, and we construct the periodogram $I_{n,i}$ on the ith block $X_{(i-1)\ell+1}, \ldots, X_{i\ell}$. The construction of this averaging of periodograms is known as Bartlett method to estimate the spectral density. Then we define our test statistic as follows:

$$Q_{n,m}(s) = \sum_{j=1}^{s} \frac{I_n(\lambda_j)}{\frac{1}{m}\sum_{i=1}^{m} I_{n,i}(\lambda_j)}. \tag{3}$$

where s is the number of Fourier frequencies we want to include in the test.
We assume that $m = m(n)$, and $\ell = \ell(n) \to \infty$ as $n \to \infty$. It is very important to emphasize the fact that m and ℓ increase with n and are not constant, and that $m = n/\ell$. We are simply using notation

m and ℓ rather than $m(n)$ and $\ell(n)$ only for the sake of simplicity. We should mention that our context is different from that of Reisen *et al* (2010), where the number of epochs is fixed.

2 Main results

In this section we give the two asymptotic distributions of $Q_{n,m}(s)$ under (a) the null hypothesis "the process has a short memory, i.e. $d = 0$" and under (b) the alternative of "the process has a long memory, i.e. $0 < d < 1/2$".
Before we state our theorem, we set its context. Let (X_t) be a linear process of the form

$$X_t = \sum_{j=0}^{\infty} a_j \epsilon_{t-j} \tag{4}$$

where (ϵ_j) are i.i.d. random variables with zero mean, variance σ^2, and finite fourth moment η, and where $\sum_{j=0}^{\infty} a_j^2 < \infty$.

Theorem 1 Let X_t be as defined in (4), $Q_{n,m}(s)$ be as defined in (3) where $s \geq 1$ is fixed and $m \to \infty$, $m = o(n)$.
(a) If X_t is a short memory process in the sense that

$$\sum_{j=1}^{\infty} |a_j| < \infty,$$

and that its spectral density f satisfies (1) with $d = 0$, then

$$Q_{n,m}(s) \xrightarrow{d} Q(s),$$

where $Q(s)$ has Gamma distribution with parameters $(s, 1)$.
(b) If X_t is a long memory process in the sense that f is of the form (1) with $0 < d < 1/2$, then

$$Q_{n,m}(s) \xrightarrow{P} \infty.$$

Comments: (a) According to the convergence result in (a), the approximated critical region of the testing procedure $d = 0$ versus $0 < d < 1/2$ at significance level α is of the form $\{Q_{n,m}(s) > \Gamma_\alpha(s)\}$ where $\Gamma_\alpha(s)$ is the quantile of order $1 - \alpha$ of Gamma distribution with parameters (s,1). This makes the test easy to implement as the critical values are available in all statistical softwares.
(b) Under the condition b) of Theorem 1, our testing procedure is consistent, i.e. the power converges to 1 under any long memory alternative $0 < d < 1/2$.

Proof of Part a) of Theorem 1. From (3) we get

$$Q_{n,m}(s) = \sum_{j=1}^{s} \frac{I_n(\lambda_j)/f(\lambda_j)}{\frac{1}{m}\sum_{i=1}^{m} I_{n,i}(\lambda_j)/f(\lambda_j)}.$$

Applying Giraitis *et al.* Th 5.3.1 (2012), we have

$$\left(\frac{I_n(\lambda_1)}{f(\lambda_1)}, \cdots, \frac{I_n(\lambda_s)}{f(\lambda_s)}\right) \xrightarrow{d} (E_1, \cdots, E_s)$$

where $E_1, \cdots, E_s$ are i.i.d. exponentially distributed with mean 1, and since clearly $Q(s) \stackrel{d}{=} E_1 + \cdots + E_s$, using Slutsky Lemma, it will then be enough to show that, for each fixed j, as $n \to \infty$,

$$\frac{1}{m} \sum_{i=1}^{m} \frac{I_{n,i}(\lambda_j)}{f(\lambda_j)} \xrightarrow{\mathrm{P}} 1,$$

or equivalently (since f is continuous, and $\lambda_j = \frac{2\pi j}{n}$, then $f(\lambda_j) \to f(0)$),

$$\frac{1}{m} \sum_{i=1}^{m} I_{n,i}(\lambda_j) \xrightarrow{\mathrm{P}} f(0) = \frac{1}{2\pi} \left(\gamma(0) + 2 \sum_{k=1}^{\infty} \gamma(k) \right). \tag{5}$$

Denoting

$$\hat{\gamma}_{n,i}(h) = \frac{1}{\ell} \sum_{k=1+(i-1)\ell}^{i\ell - |h|} X_k X_{k+h},$$

we can write

$$\begin{aligned} I_{n,i}(\lambda) &= \frac{1}{2\pi\ell} \Big| \sum_{k=1+(i-1)\ell}^{i\ell} X_k e^{i\lambda k} \Big|^2 = \frac{1}{2\pi} \sum_{s,t=1+(i-1)\ell}^{i\ell} X_s X_t e^{i(s-t)\lambda} = \frac{1}{2\pi} \sum_{|h|<\ell} e^{-ih\lambda} \hat{\gamma}_{n,i}(h) \\ &= \frac{1}{2\pi} \hat{\gamma}_{n,i}(0) + \frac{1}{\pi} \sum_{h=1}^{\ell-1} \cos(\lambda h) \hat{\gamma}_{n,i}(h). \end{aligned}$$

Hence the LHS of (5) is made of two parts. Firstly,

$$\frac{1}{2\pi m} \sum_{i=1}^{m} \hat{\gamma}_{n,i}(0) = \frac{1}{2\pi} \frac{1}{m} \frac{1}{\ell} \sum_{i=1}^{m} \sum_{k=1+(i-1)\ell}^{i\ell} X_k^2 = \frac{1}{2\pi n} \sum_{u=1}^{n} X_u^2 \stackrel{a.s.}{\to} \frac{\gamma(0)}{2\pi}, \text{ as } n \to \infty \tag{6}$$

by virtue of the ergodic theorem. Secondly, we prove the following L^2 convergence, as $n \to \infty$

$$\frac{1}{2\pi m} \sum_{i=1}^{m} \sum_{h=1}^{\ell-1} \hat{\gamma}_{n,i}(h) \cos(\lambda_j h) \stackrel{L^2}{\to} \frac{1}{\pi} \sum_{k=1}^{\infty} \gamma(k). \tag{7}$$

Observing that $(\hat{\gamma}_{n,1}(h), \ldots, \hat{\gamma}_{n,m}(h)$ are identically distributed),

$$\begin{aligned} \frac{1}{\pi} \mathbb{E} \left(\frac{1}{m} \sum_{i=1}^{m} \sum_{h=1}^{\ell-1} \hat{\gamma}_{n,i}(h) \cos(\lambda_j h) \right) &= \frac{1}{\pi} \sum_{h=1}^{\ell-1} \mathbb{E}(\hat{\gamma}_{n,1}(h)) \cos(\lambda_j h) \\ &= \frac{1}{\pi} \sum_{h=1}^{\ell-1} \gamma(h)(1 - h/\ell) \cos(\lambda_j h) \\ &\stackrel{n\to\infty}{\longrightarrow} \frac{1}{\pi} \sum_{k=1}^{\infty} \gamma(k), \end{aligned}$$

by the dominated convergence theorem, it suffices to prove that, as $n \to \infty$,

$$\mathrm{Var}\left(\frac{1}{m}\sum_{i=1}^{m}\sum_{h=1}^{\ell-1}\hat{\gamma}_{n,i}(h)\cos(\lambda_j h)\right) \to 0. \tag{8}$$

The LHS of (8) can be written as

$$\begin{aligned}
&\frac{1}{m^2}\sum_{s=1}^{m}\sum_{t=1}^{m}\mathrm{Cov}\left(\sum_{h=1}^{\ell-1}\hat{\gamma}_{n,s}(h)\cos(\lambda_j h), \sum_{h=1}^{\ell-1}\hat{\gamma}_{n,t}(h)\cos(\lambda_j h)\right)\\
&= \frac{1}{m}\sum_{u=-m}^{m}\Big(1-\frac{|u|}{m}\Big)\mathrm{Cov}\left(\sum_{h=1}^{\ell-1}\hat{\gamma}_{n,1}(h)\cos(\lambda_j h)\,,\,\sum_{h=1}^{\ell-1}\hat{\gamma}_{n,1+|u|}(h)\cos(\lambda_j h)\right)\\
&= \frac{1}{m}\sum_{u=-m}^{m}\Big(1-\frac{|u|}{m}\Big)\sum_{p=1}^{\ell-1}\sum_{q=1}^{\ell-1}\cos(p\lambda_j)\cos(q\lambda_j)\mathrm{Cov}\Big(\hat{\gamma}_{n,1}(p)\,,\,\hat{\gamma}_{n,1+|u|}(q)\Big)
\end{aligned}$$

The first equality is due to the stationarity of the process

$$\left(\sum_{h=1}^{\ell-1}\hat{\gamma}_{n,s}(h)\cos(\lambda_j h)\right)_s.$$

Therefore it will be enough to show that for $u \geq 2$,

$$\sum_{p=1}^{\ell}\sum_{q=1}^{\ell}\mathrm{Cov}\left(\hat{\gamma}_{n,1}(p), \hat{\gamma}_{n,1+u}(q)\right) \to 0, \qquad \text{as } n,\ (\text{or equivalently } \ell) \to \infty, \tag{9}$$

which is done in Lemma 1 in the appendix.

Proof of Part b) of Theorem 1. As $Q_{n,m}(s)$ is the sum of nonnegative terms, it will be enough to show that for each fixed $j = 1, \ldots, s$

$$\frac{I_n(\lambda_j)/f(\lambda_j)}{\frac{1}{m}\sum_{i=1}^{m} I_{n,i}(\lambda_j)/f(\lambda_j)} \xrightarrow{P} \infty. \tag{10}$$

Using Deo (1997), the numerator in (10) converges in distribution to a non degenerated positive random variable. Therefore it will be enough to show that

$$\frac{1}{m}\sum_{i=1}^{m}\frac{I_{n,i}(\lambda_j)}{f(\lambda_j)} \xrightarrow{P} 0 \qquad \text{as } n \to \infty. \tag{11}$$

It is well known that if $Z_n \geq 0$ and $\mathbb{E}(Z_n) \to 0$ as $n \to 0$ then $Z_n \to 0$ in probability. Therefore, since the variables involved in (11) are nonnegative, it suffices to show that

$$\frac{1}{m}\mathbb{E}\left(\sum_{i=1}^{m}\frac{I_{n,i}(\lambda_j)}{f(\lambda_j)}\right) \to 0 \qquad \text{as } n \to \infty. \tag{12}$$

Since $I_{n,1}, \ldots, I_{n,m}$ are identically distributed, the LHS of (12) is simply equal to

$$\mathbb{E}\left(\frac{I_{n,1}(\lambda_j)}{f(\lambda_j)}\right) = \frac{\ell^{-1}\sum_{s,t=1}^{\ell}\gamma(s-t)e^{i(s-t)\lambda_j}}{f(\lambda_j)} = \frac{\sum_{h=-\ell}^{\ell}\left(1-\frac{|h|}{\ell}\right)\gamma(h)\cos(\lambda_j h)}{f(\lambda_j)}. \tag{13}$$

Under (1), $f(\lambda) \sim c\lambda^{-2d}$ near zero ($0 < d < 1/2$), thus $f(\lambda_j) \sim cn^{2d}$, and the numerator above is bounded by

$$\sum_{h=-\ell}^{\ell}|\gamma(h)| \leq C_1\sum_{h=1}^{\ell}h^{2d-1} \sim C_2\ell^{2d}$$

for some positive constants c, C_1, C_2. Hence the ratio (13) is $O\left((\ell/n)^{2d}\right) \to 0$ as $n \to \infty$ since $d > 0$ and $\frac{\ell}{n} = \frac{1}{m} \to 0$.

3 Simulation Studies

In this section we illustrate the performance of our test based on $Q_{n,m}(s)$ statistic and we compare it with the well known spectral domain test proposed by Lobato and Robinson (1998) and the time domain one based on VS statistic, briefly introduced below.

Lobato Robinson test (LR test).

In a one-dimensional case, Lobato and Robinson test statistic is $LM = t^2$, where

$$t = \sqrt{m}\frac{\sum_{j=1}^{m}\nu_j(I(\lambda_j))}{\sum_{j=1}^{m}(I(\lambda_j))},$$

and

$$\nu_j = \log j - \frac{1}{m}\sum_{j=1}^{m}\log j.$$

Under certain conditions on the bandwidth m, if the spectral density is twice boundedly differentiable near 0 and $d = 0$, the statistic t converges in distribution to the standard normal distribution. Our approach is different since we use a blocking method rather than estimating $f(0)$ and we do not assume such smoothness conditions on the spectral density near zero, except the continuity of g in (1), or on the tuning parameters except that the epoch length, $\ell(n) = o(n)$. In the simulation (Figure 3) we select for m the optimal value given in Lobato and Robinson (1998).

V/S test.

This test is defined as the ratio $M_n = V_n/s_{n,q}^2$, where

$$V_n = \frac{1}{n^2}\left[\sum_{k=1}^{n}(S_k^*)^2 - \frac{1}{n}\left(\sum_{k=1}^{n}S_k^*\right)^2\right],$$

with $S_k^* = \sum_{j=1}^{k}(X_j - \bar{X}_n)$, $\hat{s}^2_{n,q} = \tilde{\gamma}_0 + 2\sum_{j=1}^{q}\left(1 - \frac{j}{q}\right)\tilde{\gamma}_j$, and where $\tilde{\gamma}$ is the sample covariance

$$\tilde{\gamma}_j = n^{-1}\sum_{i=1}^{n-j}(X_i - \bar{X}_n)(X_{i+j} - \bar{X}_n), \quad 0 \leq j < n$$

where $\bar{X}_n$ is the sample mean of $X_1, \ldots, X_n$.
The null hypothesis of V/S test consists in the short memory case (which is our case (a) in theorem 1). Under some technical conditions on the bandwidth q and the fourth order cumulants (see Giraitis *et al.* (2003) for the exact hypotheses), M_n converges to a non degenerated distribution with cdf $F_{VS}(x) = F_K(\pi\sqrt{x})$ where F_K is the asymptotic cdf of the standard Kolmogorov statistic. The authors suggest the use of $q = n^{1/3}$ for optimal results. This is the chosen value in our simulation (see Figure 3).

Choice of window parameters m and s in $Q_{n,m}(s)$ defined in (3)

The choice of m is done to ensure that the empirical size will be close to the nominal size of the test. To do this we simulate independent copies of AR(1) and we evaluate for each sample the pvalue associated to $Q_{n,m}$-test. Then we draw the empirical cumulative distribution function of pvalues. Under the null hypothesis this cdf should match the cdf of the uniform distribution.

Figure 1 represents these empirical cdfs of $Q_{n,m}(s)$ for several Gaussian AR(1) models. We take $m = n^b$, $b = .3, .4, .5, , .6$ and $n = 1000$, 5000 and 10 000. Figure 1 shows a good fitting between the line $y = x$ and the empirical cdf associated with the choice $m = \sqrt{n}$. We also note that this choice remains good showing some robustness against deviation from normality of the innovations ϵ_t (see Figure 3).

As for the parameter s, our simulation (with $m = \sqrt{n}$) studies have shown that the choice of s depends on the sample size (See Figure 2). For small to moderate sample size, a small choice of s (such as $s = 2, 3$) would be preferable. This might be due to the fact that if e.x. we take $n = 1000$, each sub-periodogram is made of $\ell = 32$ observations, and therefore only the very first Fourier frequencies (closest to frequency zero) will have a significant contribution in determining whether a data set come from a short or long memory process. We also note that when the sample size is quite large, we can still obtain a good empirical size while increasing s. Figure 3 again shows e.x., that a good choice of s would be $s \leq 5$ for $n = 5000$ and $s \leq 10$ for $n = 10,000$ We suggest that if the sample is very large, then we should slihhtly increase s.

Non Gaussian innovations

Simulations have shown that under normality assumption LR test tends to outperform all time domain tests such as V/S, KPSS, RS etc. Here we compare our $Q_{n,m}(s)$ test with LR test and V/S test under the null hypothesis when the innovations of AR(1) are non Gaussian. (we generated innovations from heavily skewed exponential and log normal distributions, but we only reported the results for exponential distribution.
Figure 3 shows a lack of robustness to non-normality of V/S and LR tests in the sense that these tests seem to suffer from high empirical size (the ratio of rejecting the null is much bigger than the nominal level $\alpha = .05$, .1 etc). On the opposite, $Q_{n,m}(s)$ test seems to have a more reasonable empirical size

(very much close to the nominal level $\alpha = .05$, $.1$ etc). Similar conclusions hold for innovations with log-normal distributions as well.

Power of $Q_{n,m}(s)$ test.

Figure 4 illustrates the behavior of the empirical power of this test when s varies. We only represent the graphs for a non-Gaussian case (innovations with log normal distribution) but there are no noticeable differences with the Gaussian case. As expected, it shows an increase in the power function with the parameter s. Since LR and V/S empirical sizes are quite high, as seen in the previous section, as compared to $Q_{n,m}(s)$ test, the former tests will tend to be more powerful than $Q_{n,m}(s)$ test especially in Gaussian case. This is very much anticipated and for this reason we did not report the power comparison graphs between V/S and LR on one side and $Q_{n,m}(s)$ on the other side.

In conclusion, our findings confirm that, mainly due to its relatively easy implementation and large sample properties, frequency-domain approach should be used more often in combination with classical time domain methods when it comes to testing for long memory or for unit root (as AR(1) coefficient tends to 1) in the time series analysis.

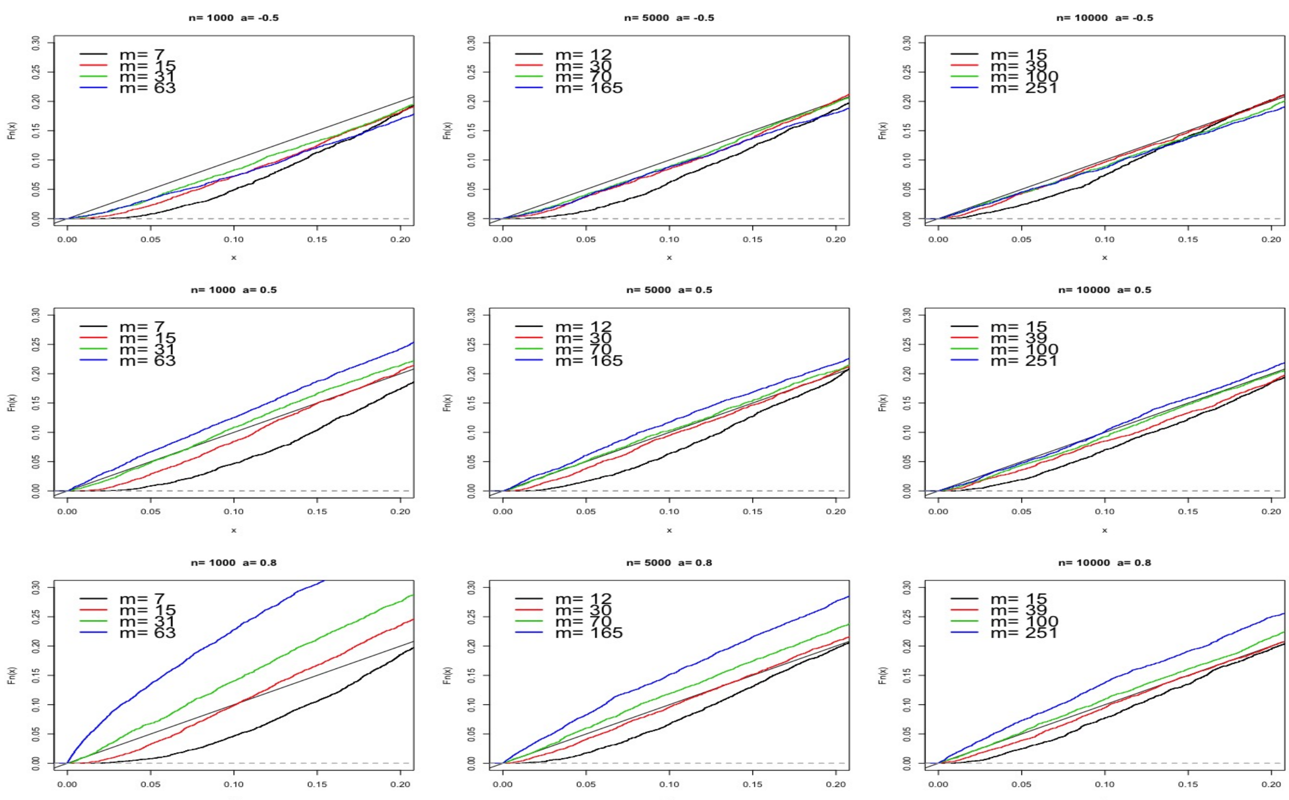

Figure 1: P-values of $Q_{n,m}(2)$; 5000 independent replications of AR(1) model with AR coefficient $a = -0.5$, 0.5, 0.8 and Gaussian innovations.

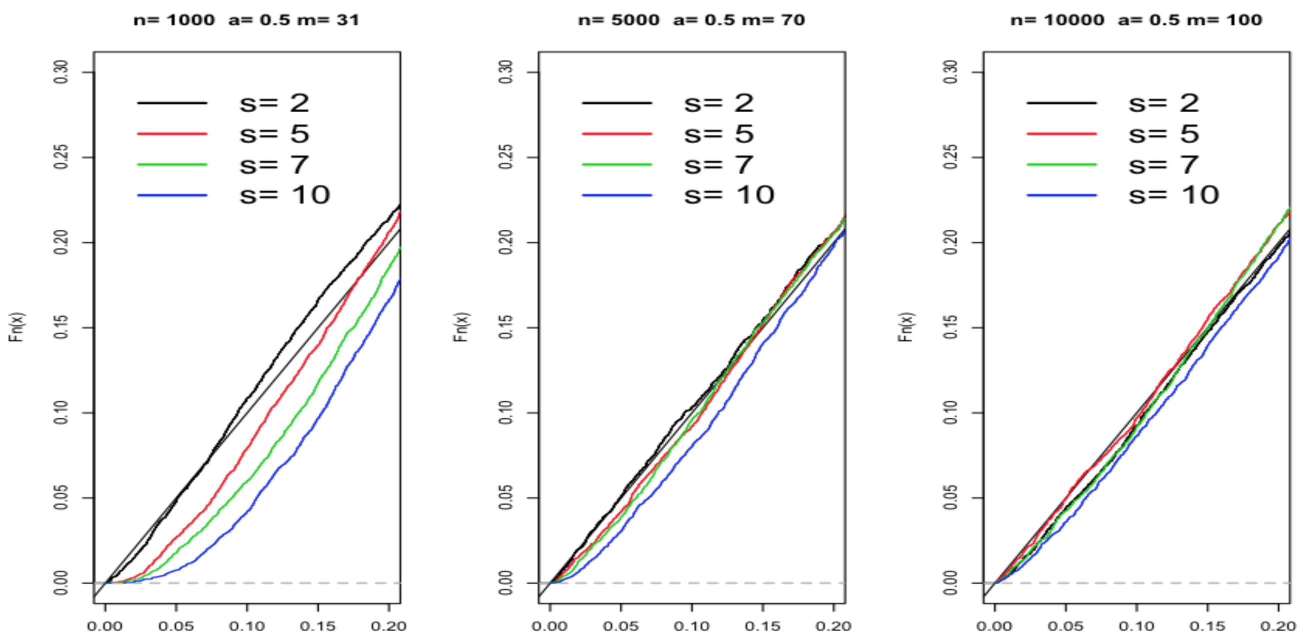

Figure 2: P-values of $Q_{n,\sqrt{n}}(s)$; 5000 independent replications of AR(1) model with AR coefficient $a = 0.5$ and Gaussian innovations.

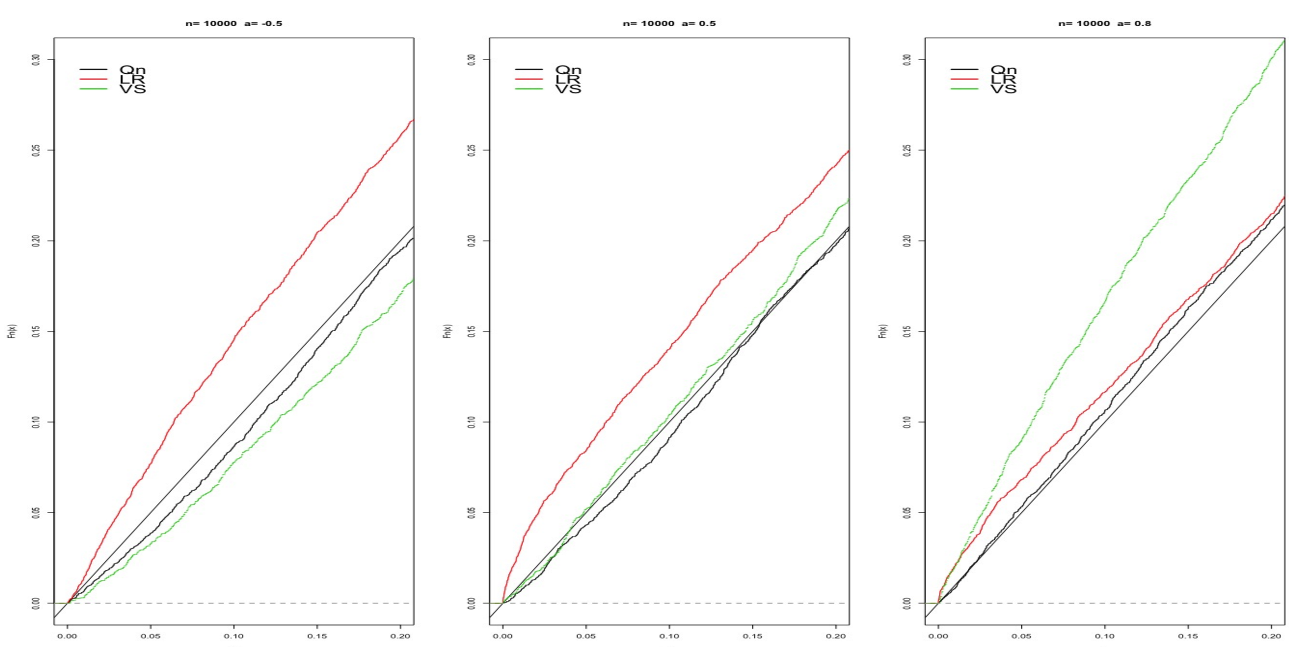

Figure 3: P-values of $Q_{n,\sqrt{n}}(2)$; 5000 independent replications of AR(1) model with AR coefficient $a = -0.5$, 0.5, 0.8 and non gaussian innovations: $\epsilon_t = \xi_t - 1$ where ξ_t are i.i.d. standard exponentials

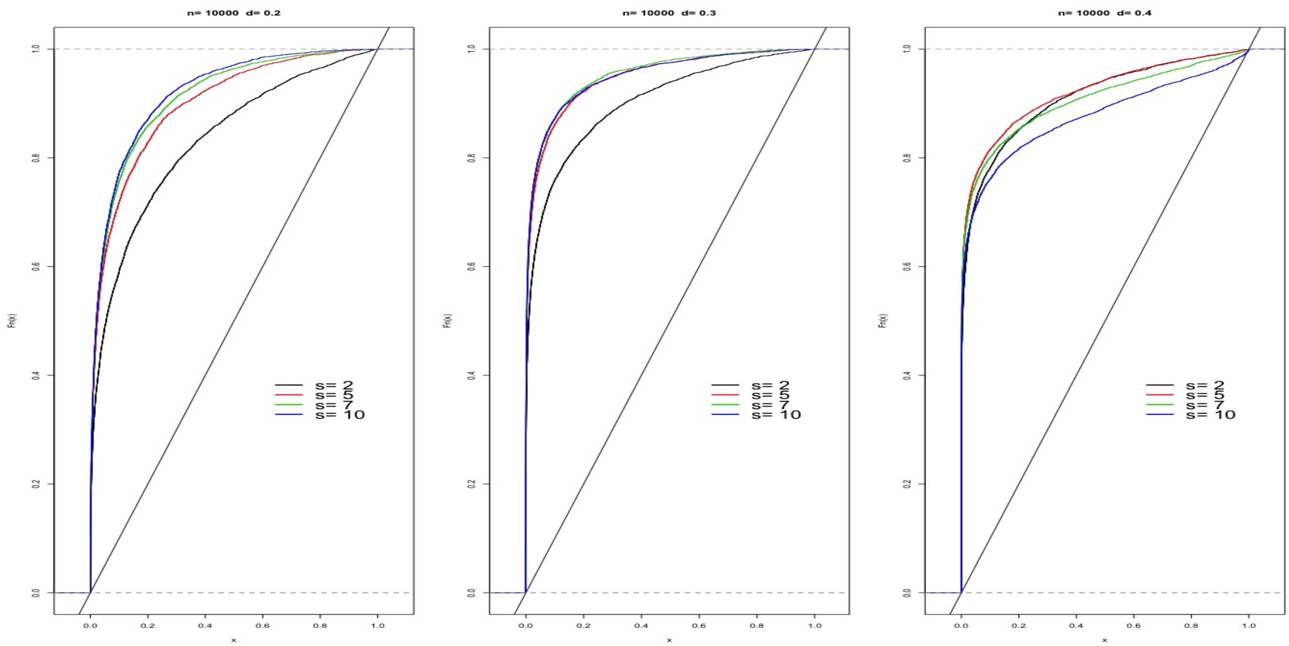

Figure 4: P-values of $Q_{n,\sqrt{n}}$(s); 5000 independent replications from long memory model ARFIMA(0,d,0): $X_t = (1-B)^d \epsilon_t$, with long memory parameter $d = 0.2$, 0.3, 0.4, and non gaussian innovations: $\epsilon_t = (\zeta_t - \sqrt{e})/\sqrt{(e-1)e}$ where ζ_t has log normal distribution, $Log\mathcal{N}(0,1)$

4 Appendix

In this appendix we state and prove the convergence (9) used in the proof of Theorem 1-a.

Lemma 1 Under the hypotheses of Theorem 1-a,

$$\sum_{p=1}^{\ell}\sum_{q=1}^{\ell} \mathrm{Cov}\left(\hat{\gamma}_{n,1}(p), \hat{\gamma}_{n,1+u}(q)\right) \to 0, \qquad \text{as } n,\ (\text{or equivalently } \ell) \to \infty.$$

Proof We recall that σ^2 is the variance of ϵ_t and η is its fourth moment. We have with the variable change $s = z - \ell u$,

$$\mathrm{Cov}\left(\hat{\gamma}_{n,1}(p), \hat{\gamma}_{n,1+u}(q)\right) = \frac{1}{\ell^2}\sum_{t=1}^{\ell-p}\sum_{z=1+\ell u}^{\ell(u+1)-q} \mathbb{E}[X_t X_{t+p} X_z X_{z+q}] - \frac{(\ell-p)(\ell-q)}{\ell^2}\gamma(p)\gamma(q)$$

$$= \frac{1}{\ell^2}\sum_{t=1}^{\ell-p}\sum_{s=1}^{\ell-q}\left[\mathbb{E}[X_t X_{t+p} X_{s+\ell u} X_{s+\ell u+q}] - \gamma(p)\gamma(q)\right]$$

$$= \frac{1}{\ell^2}\sum_{t=1}^{\ell-p}\sum_{s=1}^{\ell-q}\left[\mathbb{E}[X_t X_{t+p} X_{t+s-t+\ell u-p+p} X_{t+s-t+\ell u-p+p+q}] - \gamma(p)\gamma(q)\right].$$

Similarly to the computation used in the proof of Proposition 7.3.1 of Brockwell and Davis (1991) we then obtain,

$$
\begin{aligned}
&\mathrm{Cov}\bigg(\hat{\gamma}_{n,1}(p), \hat{\gamma}_{n,1+u}(q)\bigg) \\
&\quad = \frac{1}{\ell^2}\sum_{t=1}^{\ell-p}\sum_{s=1}^{\ell-q}\bigg[\gamma(p)\gamma(q) + \gamma(s-t+\ell u)\gamma(s-t+\ell u+q-p) \\
&\quad +\gamma(s-t+\ell u+q)\gamma(s-t+\ell u-p) \\
&\quad +(\eta-3)\sigma^4\sum_{i=0}^{\infty} a_i a_{i+p} a_{i+s-t+\ell u} a_{i+s-t+\ell u+q} - \gamma(p)\gamma(q)\bigg] \\
&= \frac{1}{\ell}\sum_{h=1+q-\ell}^{\ell-p+1}\bigg(1-\frac{|h|}{\ell}\bigg) T_{\ell,u}(h,p,q),
\end{aligned}
$$

where

$$T_{\ell,u}(h,p,q) = T^{(1)}_{\ell,u}(h,p,q) + T^{(2)}_{\ell,u}(h,p,q) + T^{(3)}_{\ell,u}(h,p,q),$$

with

$$
\begin{aligned}
T^{(1)}_{\ell,u}(h,p,q) &= \gamma(\ell u-h)\gamma(\ell u-h+q-p), \\
T^{(2)}_{\ell,u}(h,p,q) &= \gamma(\ell u-h+q)\gamma(\ell u-h-p),
\end{aligned}
$$

and

$$T^{(3)}_{\ell,u}(h,p,q) = (\eta-3)\sigma^4\sum_{i=0}^{\infty} a_i a_{i+p} a_{i+h+\ell u} a_{i+h+\ell u+q}.$$

Next for $u \geq 2$, putting $k = q-p$,

$$
\begin{aligned}
\frac{1}{\ell}\sum_{p=1}^{\ell}\sum_{q=1}^{\ell}\sum_{h=-\ell}^{\ell}\big|T^{(1)}_{\ell,u}(h,p,q)\big| &= \sum_{h=-\ell}^{\ell}\sum_{k=-\ell}^{\ell}\bigg(1-\frac{|k|}{\ell}\bigg)|\gamma(\ell u-h)\gamma(\ell u-h+k)| \\
&\leq \sum_{h=-\ell}^{\ell}\sum_{k=-\ell}^{\ell}|\gamma(\ell u-h)\gamma(\ell u-h+k)| \\
&= \sum_{h=-\ell}^{\ell}|\gamma(\ell u-h)|\bigg(\sum_{k=-\ell}^{\ell}|\gamma(\ell u-h+k)|\bigg) \\
&\leq 2\sum_{i=\ell}^{\infty}|\gamma(i)|\bigg(2\sum_{j=0}^{\infty}|\gamma(j)|\bigg) \to 0 \qquad \text{as } \ell\to\infty,
\end{aligned}
$$

since the covariances are summable in short memory. Similar computation shows also that

$$\frac{1}{\ell}\sum_{p=1}^{\ell}\sum_{q=1}^{\ell}\sum_{h=-\ell}^{\ell}\big|T^{(2)}_{\ell,u}(h,p,q)\big| \to 0, \qquad \text{as } \ell\to\infty.$$

As for the last term, permuting the order of summation, and using the variable change $s = i + h + \ell u$, and without loss of generality, considering the coefficients $a_i \geq 0$.

$$\begin{aligned}\sum_{h=-\ell}^{\ell} T_{\ell,u}^{(3)}(h,p,q) &= (\eta-3)\sigma^4 \sum_{i=0}^{\infty} a_i a_{i+p} \sum_{s=i+(u-1)\ell}^{i+(u+1)\ell} a_s a_{s+q} \\ &\leq (\eta-3)\sigma^4 \sum_{i=0}^{\infty} a_i a_{i+p} \sum_{s=\ell}^{\infty} a_s a_{s+q} = \gamma(p) o(1), \text{ as } \ell \to \infty,\end{aligned}$$

and therefore, as the covariances are summable in short memory,

$$\frac{1}{\ell} \sum_{p=1}^{\ell} \sum_{q=1}^{\ell} \sum_{h=-\ell}^{\ell} T_{\ell,u}^{(3)}(h,p,q) = \frac{o(\ell)}{\ell} \to 0, \text{ as } \ell \to \infty,$$

which completes the proof of the lemma.

References

[1] Bailey, N, Giraitis, L. (2016). Spectral approach to parameter-free unit root testing. J. Computational Statistics and Data Analysis. 100, 4-16.

[2] Beran, J., Feng, Y., Ghosh, S., Kulik, R. (2013). Long-memory processes. Monographs on Statistics and Applied Probability, (61), Springer

[3] Brockwell, P., Davis, R. (1991). *Time Series: Theory and Methods.* Springer, New York.

[4] Deo, R. S. (1997). Asymptotic theory for certain regression models with long memory errors. Journal of Time Series Analysis, 18(4), 385-393.

[5] Giraitis, L., Koul, H. L., and Surgailis, D. (2012). Large sample inference for long memory processes. London: Imperial College Press.

[6] Giraitis, L., Kokoszka, P., Leipus, R. and Teyssière, G. (2003). Rescaled variance and related tests for long memory in volatility and levels. *Journal of Econometrics*, **112**, 256–294.

[7] Giraitis, L., Leipus, R., Philippe, A. (2006). A test for stationarity versus trends and unit roots for a wide class of dependent errors. *Journal of Econometric Theory*, **22**, 989–1029.

[8] Hurst, H. (1951). Long-term storage capacity of reservoirs. *Trans. Am. Soc. Civil Eng.*, **116**, 770.

[9] Kwiatkowski, D., Phillips, P., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, **54**, 159-178.

[10] Lee D., Schmidt P. (1996). On the power of the KPSS test of stationarity against fractionally-integrated alternatives. *Journal of Econometrics*, **73**, 285–302.

[11] Lobato, I, Robinson, P. (1998). A nonparametric test for I(0). *The review of economic studies.* **65(3)**, 475-495.

[12] Mandelbrot, B., Wallis, J. (1969). Computer experiments with fractional Gaussian noises, Parts 1, 2, and 3. *Water Resources Res.*, **5(1)**, 228–267.

[13] Moulines, E. and Soulier, P. (2003). Semiparametric Spectral Estimation For Fractional Processes. Theory and Applications of Long Range Dependence. P. Doukhan, G. Oppenheim, M. Taqqu. Eds. Birkhauser.

[14] Reisen A., Moulines E, Soulier P., and Franco G. (2010). On the properties of the periodogram of a stationary long-memory process over different epochs with applications. Journal of Time Series Analysis, Vol. 31, Issue 1, pp. 20-36.